

# WHAT BELIEFS ARE NOT. INTENTIONAL REALISM WITHOUT MEANINGFUL ENTITIES IN THE BRAIN

Andreas Kemmerling

## I.

Let's call anybody an "intentional realist" if he or she claims that there are literally true sentences of the type "*X* believes (at *t*) that *p*". A "naturalist" is a person who claims that all there is, in the final analysis, belongs to the extension of those predicates which are needed for stating the basic physical laws. Is intentional realism compatible with naturalism?

If intentional properties (like '*believing that Paul is slim*') were reducible to those properties which are expressed by the predicates needed for stating the basic physical laws, then intentional realism would obviously be compatible with naturalism. But it is not obvious. In fact, we have reason to think that intentional reductionism doesn't work. I shall without further argument assume that reductionism in fact fails. Given that, how can you be an intentional realist and a naturalist as well?

I want to reject two strategies designed to show that naturalist ontology is compatible with intentional realism:

- (1) Try to find a theory which shows how to assign, in a purely physicalist way, an intentional type (e.g. "belief-that-*p*") to any belief-token.
- (2) Try to find a theory which shows how to assign, to anybody, *X*, who (at *t*) believes that *p*, a physical condition *c* such that (i) *X* is in *c* at *t*, and (ii) it is exclusively in virtue of (i) that *X* (at *t*) believes that *p*.

[Concerning (1)] The first strategy presupposes that some variant of the type/token-distinction can be applied to intentional states. This presupposition is false. Tokens are concrete particulars which exemplify abstract types. But, as I shall argue, there are no concrete things in the world which are the beliefs that it is raining; at least we have no good reason to believe in their existence. (Such things, if there were any, would be what I referred to as "meaningful entities" in the title of this paper: concrete particulars which *are* beliefs, most probably to be found in the brain.)

This whole idea of applying the type/token-distinction to intentional states is a howling misconstrual of our ordinary ways of ascribing beliefs and referring to them. Maybe this mis-

construal stems from the widespread tendency to assimilate beliefs to sentences in a language of thought; for, of course, sentences and other linguistic items are those things in the world to which the type/token-distinction applies paradigmatically. Yet the type/token-distinction does not apply to intentional states as we characteristically deal with them in our descriptions, explanations, prognoses and avowals. Before I shall try to support this claim, let me briefly sketch how our intentionalist discourse may manage to get along without bringing in intentional state tokens.

The ontology implicitly underlying our ordinary ways of using intentional idioms, if there is such an ontology, might be better captured in the following way (sticking to the example of belief): Basically there are just beliefs, believers and believings. Beliefs are abstract particulars, not universals, not types; they have no tokens. If Peter believes that it is raining and Paul believes that it is raining, they hold the same belief: the one and only belief that it is raining. Peter's holding this belief from Sunday to Monday and Paul's holding it from Tuesday to Wednesday are clearly not the same instances of holding the belief. But equally clearly, such holdings or believings cannot be construed as concrete tokens of an identical belief-type. One might want to introduce a distinction between believing-types and believing-tokens; that's fine with me. But this is a distinct distinction. Remember that believings have no truth conditions, it's only beliefs that are true or false. Remember also that nothing which could be referred to as a believing-token is related to the belief that it is raining like a token is related to a type of which it is a token. Believings, whatever they are, are not belief-tokens. And it is not clear at all, what difference to be found in ordinary intentionalist discourse is aimed at by introducing a type/token-distinction for believings. If two believings of Peter and Paul differ as to what is believed, and we want to capture this kind of difference, we can simply point to the different beliefs involved; we do not need to bring in a new category of believing-types. Of course, we may type believings according to what is believed by the believers who are undergoing their believings; this is one of the many ways in which believings can be typed. Nevertheless, the fact remains that even such believing types are not beliefs, and their tokens not 'belief-tokens'. What we have found is this: beliefs (they are abstract particulars and can be 'semantically evaluated'; but they have no tokens); believings (states of believers which cannot be said to be true or false). Introducing belief-properties makes things easier or at least more perspicuous. A 'belief-property' is what a person is said to have whenever a belief-predicate, like "— believes that it is raining", is attributed to her. Beliefs and believings may then be viewed as constructions from belief-properties. A belief is the logical offspring of any belief-property; a believing is the empirical offspring of any exemplification of a belief-property. Any well-behaved belief-predicate yields a belief; any case in which it could be correctly predicated of somebody yields a believing. This is what we find in ordinary discourse about beliefs and people believing. (It may be too much for some philosophers, but it should be enough for anybody.) At least it seems to me that we do not need anything else in order to account for this kind of discourse.

If this is right, the type/token-distinction seems not to be applicable to beliefs. But talk of belief-tokens and discussions of how best to 'type' them are, as you may know, fairly famil-

iar in certain quarters of the contemporary philosophy of mind. Therefore, even though the term "belief-token" is ill chosen anyway, it still may be interesting to find out what it is intended to refer to. I shall offer two guesses. One is that what is meant are brain-states of which a syntactical account can be given. The second guess is that what is meant are brain-states which are natural signs of the fact that the brain-owner has a certain belief.

All this recent talk of belief-tokens gets particularly disconcerting when it is assumed that one could – with regard to some particular thing, state or event  $x$  – first establish that it is a belief-token and then give (or indeed systematically refrain from giving) an 'intentional interpretation' for  $x$  (for example a 'that'-clause which gives the content of  $x$ ). This is a highly confused idea, as I shall try to argue. But maybe this idea gives us a clue as to what supporters of strategy (1) mean by the term "belief-token".

One thing is clear about what belief-tokens are supposed to be. Any entity  $x$  which deserves to be called a (human) belief-token, is supposed to be a particular brain-state or brain-event which obtains or occurs at a particular time in the brain of a particular person.<sup>1</sup> Let  $x$ , for example, be a particular neuronal situation or activity in Harvey's brain yesterday at exactly 2 p.m.; for  $x$ 's being a belief-token it is essential that  $x$  can be referred to by a certain kind of description. Such a description is usually thought of as being of the type "the brain-state which is caused by this-and-that, and which causes that-and-this". There seem to be various points of disagreement among the adherents of strategy (1) about how to fill out this form. (For example: should such a description specify only proximal causes of the brain-state, or is it O.K. to mention distal causes?) But one thing seems uncontroversial: such a description should not involve intentional terminology, (as for example the following description does: "Harvey's brain-state which is caused by his belief that all things considered, the best thing by far for him to do right now is to buy that case of *Chateau Latour 1990*, and which causes him to sign a cheque to the wine merchant"). Moreover, there seems to be disagreement about questions like the following: What is it that makes  $x$  a token of the belief that  $p$  and not a token of the belief that  $p^*$ , or about the question: What is it that makes  $x$  a belief-token and not a desire-token or any other kind of intentional-state/event/process-token?

I am not, however, concerned here with the details of such proposals. What I want to oppose is the general idea, the idea that there are concrete particulars in brains which could, in a first step, somehow be independently identified in terms of neurophysiological description and, in subsequent steps, recognized as belief-tokens and finally interpreted as tokens of the belief that  $p$ . Let me quote from Stephen Stich's *The Fragmentation of Reason*, where this idea shapes the discussion in large parts of the book.<sup>2</sup> Stich writes:

1 This seems to rule out believings. A believing-that- $p$  of a person  $P$  for a period  $t_1$ - $t_2$  is representable as  $\langle P$ , the belief that  $p$ ,  $t_1$ - $t_2$   $\rangle$ . Neurophysiologists assure me that the brain-states they are acquainted with are quite different from anything representable in this way.

2 Stephen Stich, *The Fragmentation of Reason*, MIT Press, Cambridge, Mass. 1990, pp. 103-109. *Nota bene*: This idea shapes the discussion in Stich's book, but it is not something he, at the surface, commits himself to. Yet he is committed to an idea which I take to be even worse: If this idea of beliefs as interpretable brain-state tokens doesn't work, the whole concept of belief is for the birds. That's why I take it to be fair to quote him here at length: although he doesn't believe (*sit venia verbo*) in this account, he thinks of it as a last chance for belief.

The theory I'll be using assumes that beliefs are real psychological states, not explanatory fictions like the lines in a parallelogram of forces. [...] The theory embraces the so-called token-identity hypothesis, which claims that each instance (or token) of a belief is identical with some neurophysiological state or other. [...] Subtleties aside, what these assumptions amount to is the claim that *belief-state tokens are brain-state tokens*.

Unlike most brain states, however, [...] beliefs have semantic properties. How can this be? *What is it for a brain-state token* – a neurophysiological state or happening – *to be true or false?* One familiar framework in which an answer can be developed posits the existence of a function that maps certain brain-state tokens [...] onto entities that are more naturally thought of in semantic terms, entities like propositions, or content sentences, or specifications of truth conditions. [...] An account of what it is for a belief token (i.e., a certain brain-state token) to be true can then be given in terms of the entity to which it is mapped: the belief [*sic*] is true if and only if the proposition (or content sentence) to which it is mapped is true. [...]

For argument's sake I will simply grant that some completely unproblematic story can be told about propositions and what makes them true. [...]

The idea is that beliefs are complex psychological states which, like sentences, can be viewed as built up out of simpler components. So by mapping the elements out of which beliefs are constructed to the symbols of some uninterpreted formal language, in a way that preserves well-formedness, we can associate belief tokens with well formed formulas in that language. Indeed, we can view belief tokens as neurally encoded inscriptions of the relevant well-formed formulas. *To have a belief, then, is to have a token of a well-formed formula stored appropriately in one's brain.* The question of *how beliefs get their semantic properties* can now be rephrased as a question about how we can assign truth conditions to these cerebral inscriptions. [...] Suppose that when we looked inside a person's head we saw a little box labeled "Beliefs", in which there was a large and evolving collection of what appeared to be inscriptions in a language we did not understand. The job of the interpretation function is to specify truth conditions for the sentences in the Belief Box. (Italics mine, A.K.).

(What I recommend we should conclude, if we were to find such a little box labeled "Beliefs" inside a person's head is that although Mother Nature sometimes is very funny, either her English or her philosophy sucks.) I have not quoted so extensively because I thought these citations to vividly illustrate the real problems involved in trying to combine naturalism and intentional realism. On the contrary. It would be euphemistic even to say that this whole approach starts off on the wrong foot, since it wants to start where there is no foot. I have quoted so extensively, because Stich here gives such a wonderful portrayal of the philosophical climate in which belief-token-talk flourishes.

Central to this mistaken approach is the idea that one could, as it were, first identify a belief-token (or even the complete set of belief-tokens of a particular person) and then secondly interpret it. This is a misfired analogy from cases of genuine interpretation. In genuine interpretation, we have an *interpretandum* (e.g. a text, a painting, a piece of music) which gets assigned an *interpretamentum* (e.g. another text, or a musical performance, or a dance).

In such cases, we can usually identify the *interpretandum* independently of any assumptions about a specific *interpretamentum*; and we can then, as a second step, develop an *interpretamentum*. If need arises, we can pick out one *interpretandum*-token (like e.g. a particular autograph or manuscript) and make it, so to speak, the ultimate object of interpretation – the master-*interpretandum*.

But all this does not fit intentional states in the least. Whatever beliefs are, they are not *interpretanda*.<sup>3</sup> A fortiori, they are not master-*interpretanda* either. In a list of things in the world which can be interpreted you may include texts, paintings, pieces of music, facial expressions, pieces of behaviour etc., but to rank beliefs along with such things, would be a category-mistake. Beliefs do not belong in this category. In this respect they are altogether different from sentences. Sentences and beliefs have in common that they both have semantic properties, but sentences allow for a distinction between form and content whereas beliefs do not. You can disregard the content of a given sentence and still have that sentence; you can even give a sentence a new meaning and still have (in a pretty clear sense) the same sentence. But you can't do such tricks with beliefs: If you take away the content of a belief, nothing is left, you'll have taken away the belief *simpliciter*; and any attempt at giving a belief a different content and yet keeping it would be clearly a case of trying the impossible.<sup>4</sup>

What I have just said about beliefs – that they are no *interpretanda* and that they have no intrinsic form – relies on how we normally speak of beliefs, when we make no sweeping assumptions about future findings of neurophysiology. But maybe the recent inventors of belief-tokens are not given to accepting the usual ways of speaking. I speculate that indeed one point of trying to introduce belief-tokens is that one wants to introduce a form/content-distinction for beliefs, in order to squeeze them somehow into the category of the interpretable entities.

Why would one do that? Well, the fascinating thing about interpretable entities is that you can turn your back upon their content without turning your back on them; you can push aside their semantical properties without dropping the subject. And it is often an instructive and challenging enterprise to find out how far you can get in an account of such entities without giving heed to their semantical properties. For only a syntactical approach gives you a chance of being really precise, i.e. maximally precise. So if precision is your most

3 Interpreting a belief is not what goes on when a guess is made about what somebody believes, or when a hypothesis is put forward why somebody holds a certain belief.

4 There is another relevant (and related) difference between sentences and beliefs: sentences, or at least their uses, allow for a distinction between literal meaning and other kinds of meaning, whereas beliefs just have their contents — the whole variety of the literal, the metaphorical, the ironical, the implied etc. being not applicable to them. And for that reason we have no concept of the exact and complete content of a belief, other than the concept of a belief the content of which is the same as the literal meaning of a particular sentence (as used on a particular occasion). The idea that beliefs should have an exact content is parasitic on the idea that beliefs can be perfectly expressed by sentences uttered in their literal use. Sentences have meaningful parts, so the concept of literalness is applicable to them; they have a beginning and an end, so the concept of completeness is applicable to them and their meanings. A belief neither has parts nor a beginning and an end. With regard to beliefs, speaking of exactness and completeness does not have a basis, as long as we don't connect them with entities (like sentences) which have parts and boundaries.

dearly beloved value in theories, then you will be inclined to go syntactical whenever you attempt to give a theory of anything. You then may easily feel that you don't really have a theory, as long as you could not, at least in principle, go syntactical all down the line. But then you are inevitably in need of entities which have a form, or which at least can be construed as having one. So any proof-theoretic or classically computer-oriented approach to the realm of the intentional needs meaningful entities which can be attributed a form. And that's my guess about why people would try to squeeze beliefs somehow into the category of the interpretable entities. They want computationally precise theories of the intentional, and for this purpose they stipulate intentional items which have a precise ('syntactical') form.

But maybe this theoretical aim is misguided. Precision is to be cherished only where it belongs. It belongs with theories, not with their subject-matters which sometimes turn out to be irremediably messy. And precise theories on messy subjects should precisely capture the messiness of their subjects and not indulge in some conceptual sleight of hand designed to produce the impression of a subject which is precise.

The messiness of the intentional reveals itself particularly clearly in the fact that all established principles which guide and justify our intentional ascriptions contain messy concepts and, what is more, additionally contain an essential *ceteris paribus* clause. The form of all such principles (which are part of our established practices of belief-ascription) is as follows:

If  $x$  fulfils the (messy) conditions so-&-so, then, *ceteris paribus*,  $x$  believes that  $p$ .

There is, one may say, a double messiness involved even in our best ways of principled belief-ascription. Consider: "If a normal and sufficiently competent English speaker on reflection, seriously and truthfully assents to an English sentence, and if in doing so, he says that  $p$ , then, *ceteris paribus*, he believes that  $p$ ". Is there any better way of justifying the claim that somebody believes that  $p$  than by taking recourse to such a principle? I know of none. Yet it can't be denied that concepts like 'linguistic competence', 'the English language', 'on reflection' etc. are vague; moreover, the unavoidable "*ceteris paribus*" adds a highly unspecific qualification to all this. I won't try to argue here for these general claims, but rather return to the topic of belief-states as concrete particulars which have a ('syntactic') form and which can be interpreted as having certain truth conditions.

Let me try to round up my misgivings about construing beliefs as interpretable entities and about any attempt at introducing new entities which are somehow both "belief-like" and interpretable. It's off the rails. Beliefs are essentially uninterpretable.

What can be and are interpreted, on the other hand, are belief-indicators, i.e. things which indicate that a particular person holds a certain belief. And here comes my last attempt at making sense of the term "belief-token". Maybe what "belief-token" is meant to refer to is neither believings, nor carriers of a syntactical form, but belief-indicators. There are many kinds of belief-indicators. Some are intentionally produced (like e.g. assertive utterances) and some are produced without there being an intention to indicate the belief in question. Any old doing (or refraining) can be a belief-indication of this second variety. When Harvey says

"I have caught a bad cold", then by making this assertion he expresses the belief that he has caught a bad cold. This belief – his holding this belief – also can be indicated by him unintentionally in various ways: e.g. by wearing a woolen scarf in August; by buying a certain highly expensive medicine; by refraining to go to the long awaited piano-concert for which he has tickets; etc. Furthermore, in addition to these two kinds of belief-expression, there may exist even 'natural signs' of some of Harvey's beliefs. It may be that Harvey blushes only if he believes that he has made a fool of himself. And it may be that in Harvey's brain there are plenty of yet to be discovered natural signs of what he believes. It might even turn out that Harvey's brain contains a natural sign for each and every belief he holds.

Such natural signs of beliefs would be interpretable brain states. We may first identify such a brain state of Harvey's and then ask, for example: "Does it indicate that he believes that it is raining or does it indicate that he believes that it is raining hard?". So is it this what people have in mind when they talk about belief-tokens and how to interpret them: brain-states which are natural belief-indicators? I don't think so. For natural signs usually are not identical with what they signify, nor with tokens of what they signify. (Even if Harvey's sneezing four times in a row is a completely reliable natural sign of his having a cold, the quadruplicate sneezing is neither Harvey's cold nor his cold-token.) So even if a brain-state naturally signifies a belief or a believing, this brain-state should not be identified with either of them, nor should it be called a "token" of them. And furthermore, natural signs of particular beliefs – wherever they may occur – have a content quite different from the content of the belief they are signs of. Imagine that there is in Harvey's brain a natural sign of his belief that  $p$ ; let's call this feature of Harvey's brain " $s$ ". The content of  $s$  is not the content of Harvey's belief;  $s$  is not a natural sign which signifies that  $p$ , it is a natural sign which signifies that Harvey believes that  $p$ .

Therefore, if it were in fact natural indication which is meant by the term "belief-token", then again this term would be, at best, highly unfortunate. One would better speak of "neuronal belief-indicators". But anyway, it is not quite clear why the question whether there are such things in the brain should be a central one in the philosophy of mind.<sup>5</sup> They would be just another bunch of belief-indicators, to be added to those which are already available: namely pieces of behaviour, in particular certain pieces of linguistic action.

Two remarks about believings, before I leave the topic of so-called belief-tokens. A believing is a state of a person which obtains as long as he or she holds the belief in question. The attempt at construing believings as concrete entities to be found inside the believer strikes me as being conceptually insensitive to begin with. But if, in spite of this, one tried to construe them as interpretable brain-states, then it is hard to see what else they could turn out to be but neuronal belief-indicators.

5 Neuronal belief indicators, provided there any, may have a form in virtue of which they indicate what they indicate. But the bare fact that an indicator has a relevant form does not entail that that what is indicated by it has one. A thermometer, a hygrometer and a barometer are designed to have a 'syntactical' form which is indicative of (certain aspects of) the weather, yet the weather does not have such a form.

Secondly, a believing must not be confused with the believer's intermittent states of being aware that he has the belief or with his conscious thoughts the content of which is the content of the belief. To see this, imagine the following. Harvey acquired the belief that  $p$  in 1959 and he lost it in 1972; so the relevant 'believing' had an uninterrupted existence of more than 10 years. During these years, there were occasions on which Harvey consciously thought that  $p$ , furthermore, there were occasions on which Harvey was aware of the fact that he believed that  $p$ , and finally, there were occasions on which he did things which he would not have done had he not believed that  $p$ . All these relatively short-term episodes have to do with Harvey's believing that  $p$ , but they are quite distinct kinds of mental happenings. It would be quite misleading to refer to them by phrases like "conscious belief", "occurrent belief" or "active belief" thereby suggesting that such episodes are beliefs or believings. But however you refer to them, it is easy to see that they are not Harvey's belief or believing that  $p$ , since they have clearly the wrong dates. I suspect that, in spite of this temporal discordance, some people, when they use the term "belief-token", may have in mind something of this variety: mental episodes in which Harvey is conscious of his believing that  $p$ , or other mental episodes in which the fact that Harvey believes that  $p$  is explanatory of something.<sup>6</sup>

Let me stop here with my criticism of strategy (1). In summary, what I have said so far is this. First, the concept of belief, as we know it, does not allow for a type/token-distinction. Given then that there are no such things which are, strictly speaking, tokens of beliefs, I have speculated on what could be meant by the term "belief-token" instead. I mentioned believings (or 'holdings'), belief of syntactical forms, belief-indicators and some further candidates. My second point, which I have not elaborated here, is that the whole idea of interpretable belief-like entities is hard to make sense of. On a strict reading, the first strategy presupposes a distinction which does not exist. Less strictly read, it still presupposes that there are belief-like things in the head which can be assigned the content believed. Both presuppositions seem wrong.

In brief, there seems to be no philosophically interesting sense in which the hypothesis that there are belief-tokens is feasible. But if we have no good reason to assume that there are belief-tokens, then the question of how to assign them a suitable content in a naturalist manner simply does not arise. Therefore strategy (1) collapses.

<sup>6</sup> Jerry Fodor, for example, makes a distinction between "merely dispositional" and "occurrent" beliefs (in *Psychosemantics*, MIT Press, Cambridge, Mass. 1987, p. 22), and in a later publication makes a very interesting claim on what an Intentional Realist is committed to: "...even Intentional Realists can take an instrumental view of beliefs and desires which, by assumption, aren't occurrent causes. It is only mental states that are episodes in mental processes that Realism requires Realists to be Realistic about" (B. Loewer, G. Rey (a cura di), *Meaning in Mind*, Blackwell, Cambridge, Mass. 1991, p. 317). Is a believing real only if it has an impact on the course of mental processes? Is it real only as long as it is 'active' in this manner? This would be a very uncommon view of believings and their duration.



II.

[Concerning (2)] The second strategy does not presuppose a distinction between types and tokens of intentional states. Let me repeat what it says.

(2) Try to find a theory which assigns to anybody,  $X$ , who (at  $t$ ) believes that  $p$ , a physical condition  $c$  such that (i)  $X$  is in  $c$  at  $t$ , and (ii) it is exclusively in virtue of (i) that  $X$  (at  $t$ ) believes that  $p$ .

This strategy starts from the realist assumption that persons hold beliefs at certain times, and it suggests searching for those physical factors which constitute a person's holding a particular belief at a particular time. In speaking of the intra-personal physical factors which are 'constitutive' of a person's believing that  $p$ , I am using the word "constitutive" in a fairly weak and somewhat vague sense. Strictly speaking, such factors cannot constitute a person's believing that  $p$ . At best, such factors provide what might better be called the intra-personal portion of the person's believing that  $p$ . (But let us here leave aside the problems arising from the fact that intentional content depends on extra-individual features). The adherent of strategy (2) hopes that starting from intentional facts (like the fact that Harvey believes that it is raining) he may be able to isolate the 'relevant' physical features within Harvey's body-limits. Indeed, he may even hope that if all goes well, and those internal factors turn out to have a clear-cut neurophysiological structure, then he would have discovered something which may well lead to a legitimate type/token-distinction for intentional states like believing-that- $p$ .<sup>7</sup>

This strategy is much better than the first one, since it doesn't rest on a confused ontology of the intentional. But alas, it is factually hopeless. An adherent of this approach sets out to search for the 'relevant' features of the believer which can be considered as constitutive of his believing that  $p$ . So let us ask: what is constitutive of somebody's believing that  $p$ ? The best clue we have is what we recognize as crucially relevant in ascribing the belief that  $p$ , or in testing the claim that a particular person believes that  $p$ .

What are those factors which are relevantly involved in literally true belief-ascriptions? It may be a good idea to restrict ourselves at present to the clearest cases we can provide, i.e. cases where it is obviously true that a particular belief-predicate applies to somebody. Consider an example of such a case. Harvey has just said "It's raining"; he is a competent speaker of English, he was serious in uttering this sentence, he was paying full attention to what he was

7 Note that this is not the distinction on which strategy 1 depends. The adherents of strategy 1 want to find a way of assigning intentional content to brain-state-tokens; whereas the followers of strategy (2) have it the other way round: they want to assign to any belief a person holds at a time a neurophysiological structure which is present in that person's brain at this time. The brain-state  $b$  at which they might arrive is not up for intentional grabs; it's not an entity the intentional interpretation of which is an open question. Within the framework of strategy (2), the only thing that makes  $b$  a believing-that- $p$  is the fact that the owner of the brain in question believes that  $p$  and that  $b$  is the collection of those factors which are the internal physiological correlates of the fact that the brain-owner has that belief.

saying, he was not attempting to mislead us about the weather, and so on. Here we have a most obvious case in which we can literally say: "Harvey believes that it is raining". Which factors of the situation make this a most obvious case of a literally true belief-ascription? Well, let me mention a few: linguistic competence, seriousness, truthfulness, attention to what one is saying. Do we have any hopes for a naturalist account of these factors? I don't think we should. The chances of somebody's developing (within the next ten generations) a theory which isolates the physical factors which constitute serious assertion look mighty slim.

Some philosophers feel that the question whether or not we can hope for a naturalist account of serious assertion is beside the point, for serious assertion has to do only with belief-ascription, not with belief itself. A somewhat more elaborate version of this objection goes like this: "In order to give a naturalist account of what Harvey's believing that it is raining consists in, one does not have to give a naturalist account of those factors which constitute the best evidence of its being the case that Harvey has that belief. In fact, suggesting such an intimate connection between intentional facts (like the fact that Harvey believes that it is raining) and evidence of such facts smells of anti-realism with regard to beliefs – and so misses the point of strategy (2), which is to save intentional realism."

Let me try to answer this objection. First, it has to be noted that the Intentional Realism defended here is committed only to the claim that there are literally true sentences of the type "X believes that *p*". This modest kind of Intentional Realism is silent about issues concerning the relationship between facts and evidence. Moreover, it does not entail any claims on the causal efficacy of intentional states. What my brand of Intentional Realism says is simply that sometimes somebody *really* believes something. It is sometimes literally true that somebody believes something. Among the cases where predicates of the type "believes that so-and-so" are ascribed, there are many cases where they apply only metaphorically or in some other deficient manner; but there are cases in which they apply strictly and straightforwardly. In this respect, there is nothing fishy or second-class about them.

It has, however, to be conceded that in many cases it is not completely clear if such predicates are applied strictly or only as a manner of speaking. When belief-predicates are applied to thermostats, chess-computers, bodily organs and dolphins, what is said should not be taken literally because these are – and are known to be – not genuine believers. And when the belief that Socrates is a neurotic is attributed to Alcibiades, again what is said cannot be literally true because the ascription involves a conceptual anachronism. What is required for a belief-ascription to be true *sensu stricto* is at least a *genuine* believer and a belief-predicate which captures specifically something he or she believes.

So when you are looking for intentional facts which are characterizable by sentences of the "X believes that *p*"-variety, you should make sure that "X" refers to a genuine believer (the literalness-condition) and that the very words "*p*" capture something which X specifically believes (the exactness-condition). If one of these conditions is not fulfilled, the belief-sentence is not true *sensu stricto*, and you have failed to characterize an intentional fact.

But it is hard to draw a precise line between those cases where the ascription of an intentional property is literally true and the colourful variety of cases where such an ascription is

not quite correct. To avoid futile controversy on this topic, it is recommendable for an adherent of strategy (2) to focus his or her attention on those cases where the application of an intentional predicate may yield a statement which is literally true. This leads us to normal human beings who are competent speakers of a natural language.

But even with regard to such people two kinds of cases have to be distinguished. There are cases where we can with good reason ascribe one specific belief-predicate and make strong claims on what exactly is believed. On the other hand there are cases where there is not one particular predicate which stands out markedly as the one to be ascribed, but where a whole bunch of clearly non-synonymous predicates could be applied with equally good reason.

I am not thinking here of philosophical niceties such as radical interpretation or translation. My point concerns the mundane difference between telling what a person believes when (a) she has told us what she believes and (b) when she has not told us, although we may have other behavioural evidence as rich as you like. In situations of the first kind, on the assumption that everything relevant is normal, the class of justifiably applicable belief-predicates shrinks down to those which would be delivered by correct interpretations of what was said in the speaker's utterance. In situations of the second kind, much less is excluded. Let us leave aside issues concerning translational and interpretational indeterminacy. In situations of type (a) we then will be able to specify exactly what is believed. In these situations, there will be just one belief-predicate which stands out: When the subject asserts the sentence "It's raining", the predicate "believes that it's raining" stands out uniquely. Whereas in situations of type (b), there always will be indefinitely many intentional predicates which differ in meaning but can be applied with equally good reason. As a basis for identifying one predicate which specifies "the very belief in question" or "exactly what is believed", non-linguistic behaviour is simply not sufficient. This is so because arbitrarily fine semantical distinctions apply. The belief that it's raining is not the same as the belief that there is a considerable amount of water falling down from the skies. The belief that Harvey is very, very sweet is not the same as the belief that Harvey is very, very, very sweet.

One may think of these distinctions as all too subtle. But as soon as one neglects them, one moves onto slippery ground. For our only criterion for distinguishing beliefs is this: Could somebody seriously claim that he believes that so-and-so but not such-and-such? And it is quite clear from what we know about what people would get away with in front of a judge and jury, that this criterion leads to an extremely fine-grained individuation of beliefs. Non-linguistic behaviour simply is not delicate enough to live up to these demanding standards of belief-individuation. And as far as I can see, we have no other standards at hand.

For better or worse, sometimes some people say some of what they think. It is here, within this practice, where our common concept of belief is at home. The best clue – and in fact, I think, the only clue – we have, if we are looking for faultless cases of belief ascription (with genuine believers and belief indicating devices with 'strict contents'<sup>8</sup>), is when competent

8 For want of a better expression, I shall say that an item  $x$ , e.g. a sentence as used on a particular occasion, has  $c$  as its strict content if " $x$  means  $c$ " is unqualifiedly true, and for every  $c^*$  (where  $c^*$  and  $c$  are not exactly the

speakers seriously make a statement or perform some other assertoric speech act. In doing this, these speakers do something which has the right kind of semantical granularity for specifying beliefs *precisely*, because we can turn to the strict content of what they have said. In a sentence which specifies the strict content of the assertion made by the speaker (or in any accurate translation of this sentence) there are as many distinguishable meaningful parts as in the that-clause of a belief-predicate which satisfies the exactness-condition. This is particularly salient in the paradigmatic case, where a speaker of our language asserts an eternal sentence, because the sentence uttered itself provides us with the that-clause of the belief-predicate.

My point is this. There is only one class of trouble-free cases in which the literal-condition and the exactness-condition are fulfilled: namely when a speaker explicitly says something he or she believes.<sup>9</sup> The relevant factors involved in these cases – the factors which make these cases first class – are the ones I have just mentioned: linguistic competence, seriousness, normality and so on. Take any of them away, and you do not have a first class case of belief-ascription any more. Hence I find it hard to see how these factors could be omitted in an account which attempts at specifying the intra-personal factors constitutive of believing-that-*p*. And therefore it seems that a theory of the type which strategy (2) aims at, cannot avoid trying to give a naturalist account of these factors.

And this is why strategy (2) is hopeless. It aims at a theory which, as a matter of fact, cannot be delivered in the foreseeable future, although it may not be conceptually incoherent or ontologically confused to dream of it.

### III.

Many of us started out with type/type reductionist hopes, because we felt that this was the most straightforward way of being naturalists. For various reasons, these hopes have diminished. So we tended to look for something close to type/type reductionism in order to keep the naturalist faith. The two strategies which I have just sketched are in this spirit of trying for 'the next best thing'. But even these next best things are no good, it seems.

So let us ask: What does a naturalist need? Basically, I think what he needs is an overall worldview without any non-physical entities (in particular without any non-physical con-

same content) it holds that "*x* means *c*" is not unqualifiedly true.

<sup>9</sup> In fact, I take these cases to be *conceptually* central, because I think that "*X* believes that so-and-so" can be interestingly paraphrased by something like the following: "If *X* were a normal competent speaker of our language, then, *ceteris paribus*: *X* would – if asked whether so-and-so – answer with unqualified affirmation, therein being serious, reflective and truthful". If the belief-sentence is true, the paraphrase is true, and *vice versa*; if the paraphrase is false, the belief-sentence is false, and *vice versa*. And what is more, the two sentences seem to covary with regard to cases in which they sound odd. The paraphrase sounds odd, whenever a value is chosen for "*X*", such as, e.g., a dog or a computer, of which it is not clear what its being a normal competent speaker of our language is supposed to amount to; in these cases, the belief sentence sounds odd too. I don't want to argue here for this claim. But whatever your account of the meaning of belief-ascribing sentences, you should consider the cases in which a speaker explicitly says what she believes as particularly uncontroversial cases.

crete particulars). Borrowing a distinction from Quine,<sup>10</sup> we might say that for the naturalist ontology comes first and ideology comes second. That is to say, the naturalist is after a physicalistically clean ontology in the first place, and he cares about the ideology of his worldview only as a *cura posterior*. The ontology of a worldview concerns the question of what entities are quantified over when it is paraphrased in the notation of (objectual) quantification. The ideology concerns the predicates used in such a paraphrase. The naturalist's *cura prima* is to make sure that all (concrete) entities which are quantified over are perspicuously physical. It is here where we should stick to our reductionist ambitions. The naturalist's *cura posterior* is to operate with nothing but physicalistically clean predicates. It is here where we should abandon our reductionist reveries. As enlightened naturalists, we can allow for any kinds of predicates as long as their extensions pass naturalist muster. So if the basic intentionalist terminology consists of predicates which apply to physical systems, that's good enough. Be a naturalist about the ontology, but when it comes to ideology, feel free to be an intentional realist. That's what I recommend.

In this spirit, I shall suggest, in a very vague and sketchy manner, a much less ambitious strategy for the intentional realist who has a penchant for naturalist ontology:

(3) Give up the hope of isolating a well-defined particular physical trait  $x$  in a subject who (at a particular time  $t$ ) believes that  $p$ , such that  $x$  could be reasonably considered as the subject's belief-state. Instead try to give an account of the fact that we find it quite natural (and even inevitable) to make certain of our predictions and explanations in intentional terms.<sup>11</sup>

[Concerning 3] The idea is this: Instead of searching for meaningful entities in the brain, search for something quite different, namely a plausible explanatory story about why creatures like us have developed the sophisticated machinery of intentional discourse and why it is such a success. You have naturalized intentionality as soon as you have answered the question: "What is the adaptive value of intentional discourse?" If it turns out that it doesn't have adaptive value, you should be able to tell a convincing story about why creatures like us got hooked onto intentional discourse in spite of its pointlessness.

This may seem too weak. It may be feared that giving up hope of isolating meaningful entities in the brain is tantamount to giving up intentional realism in a naturalist framework. But meaningful entities in the brain are not needed for intentional realism; intentional facts about persons will do. A sentence like "Harvey believed at 6.00 CET that it is raining" may be literally true, even if there is no sharply outlined region of the physical world which could be reasonably regarded as Harvey's belief or believing. Therefore it may well be that, ontolog-

10 See Willard V. Quine, *Ontology and Ideology*, "Philosophical Studies" 2 (1951), pp. 11-15.

11 Wittgenstein's later writings can be read as pointing in this direction, but probably he would have been pessimistic about the possibility of "giving an account" for the fact that we find it quite natural to use intentionalist terminology.

ically speaking, there are no such entities as beliefs or believings in the brain. But a world with no such entities may be still a world with intentional facts: facts about the intentional properties of physical objects or systems. As a naturalist you refuse to allow for concrete particulars which cannot be reduced to, or decomposed into, concrete particulars which belong to the extension of those predicates which are needed for stating the basic physical laws. And you want the physical facts to be basic among the contingent facts. But you may allow for non-physical facts (i.e. facts which cannot be described in purely physical terminology), as long as they are perspicuously dependent upon physical facts and require no extra-particulars.

And indeed there is a pretty good sense in which intentional facts about persons are dependent upon physical facts and require no extra-particulars over and above what naturalist ontology provides us with. Take a complete description of all physical facts; a demiurge who builds a new world according to this description (and brings in no additional entities), cannot but create a world which contains the same intentional facts. But take a complete description of all intentional facts; a demiurge who builds a new world according to this description can come up with a world which contains different physical facts. This asymmetry between physical and intentional facts, also known as 'global supervenience', provides a good sense for saying that intentional facts are dependent upon physical ones. So what I want to suggest, with regard to the relationship between the physical and the intentional, is that we settle for global supervenience.

Let me close with a remark on intentional properties and intentional facts. What good, you may wonder, is the avoidance of concrete intentional particulars (like neuronal belief-tokens in the brain) if we are still embarrassed by intentional properties and facts? Are we therefore not still left with lots of intentional entities which defy physicalist description? Yes, we are left with them, but then again, no, we should not feel embarrassed by them. For, following Stephen Schiffer's recent suggestions,<sup>12</sup> we may treat intentional properties and facts, like all properties and facts, as pleonastic: as abstract entities derivable from manners of speaking which carry no genuine ontological commitments with regard to concrete particulars. And taking up Quine's distinction between ideology and ontology, we may find comfort in the idea that the only concrete particulars in our ontology are physical objects. For if the essence of naturalism is, in a phrase borrowed from Quine, "a repudiation of mind as a second substance, over and above body",<sup>13</sup> we can be naturalists and intentional realists at the same time: without assuming meaningful entities in the brain.<sup>14</sup>

12 See Stephen Schiffer, *Remnants of Meaning*, MIT Press, Cambridge, Mass 1987, cap. 6; and his *Language-Created Language-Independent Entities*, "Philosophical Topics", 24 (1996), pp. 149-167.

13 Willard V. Quine, *Quiddities*, Harvard UP, Cambridge, Mass. 1987, p. 133.

14 During the last years, I have read various versions of this paper to various audiences. Thanks for comments are due to Arthur Collins, Michael DePaul, Gary Ebbs, Frances Egan, Jerry Fodor, Gary Hatfield, Wolfgang Künne, Brian Loar, Brian MacLaughlin, Colin McGinn, Felix Mühlhölzer, Tobias Rosefeldt, Marina Sbisà, Stephen Schiffer and John Searle.